CIRRELT

Centre interuniversitaire de recherche
sur les réseaux d'entreprise, la logistique et le transport

Interuniversity Research Centre
on Enterprise Networks, Logistics and Transportation

Hub Network Design Problems with Profits

Armaghan Alibeyg
Ivan Contreras
Elena Fernández

June 2015

CIRRELT-2015-19

Bureaux de Montréal :
Université de Montréal
Pavillon André-Aisenstadt
C.P. 6128, succursale Centre-ville
Montréal (Québec)
Canada H3C 3J7
Téléphone : 514 343-7575
Télécopie : 514 343-7121

Bureaux de Québec :
Université Laval
Pavillon Palasis-Prince
2325, de la Terrasse, bureau 2642
Québec (Québec)
Canada G1V 0A6
Téléphone : 418 656-2073
Télécopie : 418 656-2624

www.cirrelt.ca

UNIVERSITÉ LAVAL · McGill · Concordia · ÉTS · UQÀM · HEC MONTRÉAL · POLYTECHNIQUE MONTRÉAL · Université de Montréal

Hub Network Design Problems with Profits

Armaghan Alibeyg[1,*], Ivan Contreras[1], Elena Fernández[2]

[1] Interuniversity Research Centre on Enterprise Networks, Logistics and Transportation (CIRRELT) and Department of Mechanical and Industrial Engineering, Concordia University, 1515 St-Catherine Street W., EV4 139, Montréal, Canada H3G 1M8

[2] Department of Statistics and Operations Research, Technical University of Catalonia, Barcelona, Campus Nord UPC, C5 building, level 2, room 208, C/ Jordi Girona 1-3, 08034 Barcelona Spain

Abstract. This paper presents a class of hub network design problems with profit oriented objectives. Potential applications arise in the design of air and ground transportation networks, where companies need to jointly determine the location of hub facilities as well as the design of the hub network. For this, strategic network design decisions must be integrated within the decision-making process. Such decisions may include the selection of the origin/destination nodes that will be served as well as the activation of different types of edges. This class of problems considers the simultaneous optimization of the collected profit, the setup cost of the hub network and the transportation cost. Several alternative models that can be used in a variety of situations are proposed and analyzed. For each model an integer programming formulation is presented and computationally tested in terms of both, the structure of its solution network and the difficulty for solving it with a commercial solver.

Keywords. Hub network design, hub location, profits, discrete location.

Acknowledgments. This research was partially funded by the Spanish Ministry of Science and Education (Grant MTM2012-36163-C06-05 and ERDF funds). The research of the first two authors was partially funded by the Natural Sciences and Engineering Research Council of Canada (NSERC) under grants 418609-2012. This support is gratefully acknowledged.

Results and views expressed in this publication are the sole responsibility of the authors and do not necessarily reflect those of CIRRELT.

Les résultats et opinions contenus dans cette publication ne reflètent pas nécessairement la position du CIRRELT et n'engagent pas sa responsabilité.

---

* Corresponding author: Armaghan.Alibeyg@cirrelt.ca

Dépôt légal – Bibliothèque et Archives nationales du Québec
Bibliothèque et Archives Canada, 2015

© Alibeyg, Contreras, Femández and CIRRELT, 2015

1. Introduction

Hub-and-spoke networks are frequently employed in transportation and telecommunication systems to efficiently route commodities between many origins and destinations. One of the key features of these networks is that direct connections between origin/destination (O/D) pairs can be replaced by fewer, indirect but privileged connections by using transshipment, consolidation, or sorting points, called *hub facilities*. This reduces the total setup cost at the expenses of increasing some individual transportation costs. Overall transportation costs may also decrease due to the bundling or consolidating of flows through inter-hub arcs.

Hub Location Problems (HLPs) deal with joint location and network design decisions so as to optimize a cost-based (or service-based) objective. The location decision focuses on the selection of a set of nodes to place hub facilities, whereas the network design decisions deal with the selection of the links to connect origins and destinations, possibly via hubs, as well as the routing of commodities through the network. Typically, HLPs assume that hubs must be located at the nodes of a given network, distances satisfy the triangle inequality, and there is a constant discount factor on the transportation costs of the arcs connecting hubs. In addition, classical HLPs impose that all flows are routed via the selected hubs and ignore all arc setup costs. Such problems have optimal solutions where an arc exists connecting each pair of hubs, so optimal routing paths consists of at most three arcs, two arcs connecting non-hub nodes and hub nodes, plus one intermediate arc connecting two hub nodes. This optimality condition implies that the network design decisions are mainly determined by the allocation of non-hub nodes to hubs (see, Contreras, 2015), and has been extensively exploited to develop formulations and solution algorithms for solving these classical HLPs.

Hub Arc Location Problems (HALPs) no longer assume that the above optimality condition holds, and incorporates explicit decisions on the arcs connecting two hub nodes that can be used. HALPs, in which setup costs for such arcs may be considered, were introduced in Campbell et al. (2005) and further studied in Contreras and Fernández (2014). Other HALPs impose particular topological structures on the network induced by the solution, such as tree-star (Contreras et al., 2010), star-star (Labbé and Yaman, 2008), ring-star (Contreras et al., 2015), and hub lines (Martins de Sá et al., 2015; Martins de Sá et al., 2015).

In most hub location applications arising in the design of distribution

and transportation systems, a profit is obtained for serving (i.e. routing) the demand of a given commodity. Capturing such profit may incur not only a routing cost but also additional setup costs, as the O/D nodes of the commodity may require the a priori installation of transport infrastructure. Classical HLPs and HALPs, however, ignore such profits and associated setup costs, as reflected by the requirement that the demand of every commodity must be served. Indeed, the overall profit obtained when all the commodities must be served is constant, and it does not affect the optimization of the distribution system. Broadly speaking, this requirement expresses the implicit hypothesis that the overall costs of solution networks will be compensated by the overall profits. Of course, such hypothesis does not necessarily hold, and incorporating decisions on the O/D nodes that should be served and their associated commodities may have important implications in the strategic and operational costs. To the best of our knowledge, the study of HLPs that incorporate explicit decisions on the nodes to be served has not yet been addressed in the literature.

The main focus of this paper is the study of HALPs that integrate within the decision-making process additional strategic decisions on the nodes and the commodities that have to be served. For this we introduce *Hub Network Design Problems with Profits* (HNDPPs), in which such type of decisions are incorporated within alternative modeling frameworks. HNDPPs consider a profit-oriented objective which measures the tradeoff between the profit of the commodities that are served and the overall network design and transportation costs. HNDPPs generalize HLPs and HALPs as they incorporate one additional level to the decision-making process.

We work under very mild modeling assumptions. First, we do not impose a predefined topology to solution networks. Instead, we consider setup costs for all types of edges to allow the model to determine the optimal network structure. As a result, the hub-level network may consist of more than one connected component and could even have isolated hub nodes. Second, we allow the use of edges connecting two hub nodes without a discount factor on the transportation costs (see, Campbell et al., 2005). Third, following the above comments we impose that the O/D nodes of the commodities that are routed must be activated (*served*), although a commodity may not be routed even if its associated O/D nodes are served. As a result, we consider setup costs not only for the hub facilities but also for the served nodes. HNDPPs focus on the following strategic decisions: *i)* what commodities to serve (this also dictates the nodes to activate); *ii)* where to locate the hubs; and, *iii)*

what edges to activate and of what type. As usual, the operational decisions determine how to route the commodities that are selected to be served.

Potential transportation applications of HNDPPs arise in the airline and Less-than-Truckload industries. As an example, in the case of airline companies network planners have to design their transportation network when they are first entering into the market, or may have to modify already established hub-and-spoke networks through alliances, merges and acquisitions of companies. The involved decisions are to determine the cities that will be part of their network, i.e. what cities they will provide service to (served nodes) and what O/D flights to activate (served commodities) in order to offer air travel services to passengers (served demand) between city pairs. Additional decisions focus on the location of their main airports (hub facilities) and on the selection of the legs used for connecting regional airports (served nodes) with hub airports and for connecting some hub airports between them. Finally, the transportation of passengers using one or more O/D paths on their established network. The objective is to find an optimal hub network structure that maximizes the total net profit for providing air travel services to a set of O/D flights while taking into account the (re)design cost of the network. Depending on the regulations or the company service policy, passenger air travel services could be provided: *i)* only to city pairs that are profitable, *ii)* between all city pairs that are served by the company, or *iii)* to a percentage of them (private companies with service commitment or with market penetration policies).

In this work we introduce HNDPPs and propose several alternative models of increasing complexity, which can be suitable for different application frameworks. Each model is analyzed and a mathematical programming formulation is presented and computationally tested in terms of both, the structure of the solution networks it produces and its difficulty for being optimally solved with a commercial solver. We start with two solely profit-driven HNDPPs, in which the criterion that guides decisions is profit. Such models are applicable to private companies where their ultimate goal is to maximize their net profit, independently of any other consideration. According to these models, the company would only provide service to O/D nodes that increase its profit. In the first model, among all commodities associated with served O/D nodes, only the profitable ones are routed. The second model takes into account that, due to governmental regulations or to some other reasons, a company may be forced to route commodities between served O/D nodes even if this would reduce its profit. That is, the model forces to

provide transportation services to any commodity where both its origin and destination nodes are served. We then study models that guarantee a minimum service level. This threshold is imposed through a percentage of the total number of O/D pairs (i.e commodities) that are served in one case, and through a percentage of the total demand of commodities (i.e amount of flow) in a second variation. Finally, our last models extend the previous ones by considering that for each O/D pair, the amount of commodity that requires transportation services depends on the selected profit level (among a discrete set of possible levels). These models provide a more realistic modeling framework for designing hub networks with a profit-oriented objective, at the expense of considerable increasing the complexity for solving them with a general purpose solver.

The remainder of the paper is organized as follows. Section 2 reviews the most relevant literature related to HNDPPs. Section 3 gives the formal definition of HNDPPs, whereas the introduction of the alternative types of models together with their mathematical programming formulations are given in Section 4. Section 5 describes the computational experiments we have run. The results produced by each model are presented and analyzed. The results of the different models are compared among them. The paper ends in Section 6 with some comments and conclusions.

2. Literature Review

HNDPPs are related to two families of HLPs: Hub Covering Problems (HCPs), and Competitive Hub Location Problems (CHLPs). HCPs impose that commodities between O/D pairs have to be delivered within a time limit (service level). It is implicitly assumed that a commodity is served whenever its O/D nodes are within a predefined radius of some hub node. Because HCPs restrict the length of the arcs of O/D paths to a given coverage radius, applications of these problems frequently arise in the design of telecommunication networks, where the signal deterioration must be taken into account (Campbell and O'Kelly, 2012; Yildiz and Karasan, 2015). Campbell (1994) introduces different HCPs, which have also been studied and extended by other authors (see Alumur and Kara, 2008, for a review on HCPs). More recently, Hwang and Lee (2012) study the uncapacitated single allocation p-hub maximal covering problem, which maximizes the overall demand that can be covered by p facilities within a fixed coverage radius. HCPs usually focus solely on service objectives and ignore completely transportation costs.

We are only aware of the work of Lowe and Sim (2012) that studies a HCP that considers jointly hubs setup costs and flow transportation costs, subject to covering constraints, and the work of Campbell (2013) that presents a continuous approximation model that seeks to minimize transportation costs while imposing high levels of service that discourage the use of circuitous routings. Similarly to HNDPPs, in many HCPs it is possible that some commodities remain unserved. However, in contrast to HNDPPs, HCPs implicitly assume that the setup cost for providing service to all O/D nodes is zero and thus, they do not incorporate decisions on the nodes to be served.

While most HLPs are concerned with the design of the hub network of a single firm, CHLPs consider the design of hub networks within an environment in which several firms exist in a market and compete to provide service to customers. In CHLPs each commodity chooses the competing firm that will serve its demand, based on several criteria such as travel time or service cost. The usual objective is to maximize the market share of some firm in this competing environment. Marianov et al. (1999) introduce CHLPs with two competitors in which the follower looks for the best location for a set of hubs so as to maximize the captured demand. The first model assumes that a commodity demand will be fully captured if its routing cost does not exceed the current competitor's cost. A more realistic model is also considered, in which the proportion of the commodity demand that is captured is modeled using a stepwise linear function, which is used for the comparison with the competitor's routing costs. In both models, at most one path can be used to route commodities between each O/D pair. Eiselt and Marianov (2009) extend these models to allow using more than one path to connect an O/D pair. The proportion of commodity demand that is routed on a particular path is modeled with a gravity-like attraction function that depends on both, the routing cost and the travel time.

Gelareh et al. (2010) present a model arising in liner shipping networks, where a new liner service provider designs its network to maximize its market share, using a stepwise attraction function, which depends on service times and routing costs. Lüer-Villagra and Marianov (2013) study a competitive model in which a new company wants to enter the market of an existing company. The aim is to determine the prices to charge to served commodities so as to maximize the profit of the entering company, rather than its market share. Commodities preferences for the selected firm and service route are modeled using a logit model. O'Kelly et al. (2014) present a model with price-sensitive demands. It considers three different service levels for routing

commodities between O/D pairs that use either two-hub O/D paths, one-hub O/D paths or direct connections. The model is formulated as an economic equilibrium problem that maximizes a nonlinear concave utility function minus the routing costs and the setup cost for the location of the hubs.

CHLPs have also been studied under a game theoretic framework, such as Stackelberg hub location models, cooperative game theoretic models with alliances and mergers, and non-cooperative game theoretic models (see Adler and Smilowitz, 2007; Lin and Lee, 2010; Asgari et al., 2013; Sasaki et al., 2014; Contreras, 2015). We note that HNDPPs can be clearly differentiated from CHLPs, as the focus of the former is to optimize the individual decision related to one single firm rather than on competition aspects. Furthermore, to the best of our knowledge, besides Sasaki et al. (2014) all CHLPs that have been previously studied focus on the location of the hub facilities and do not explicitly consider the selection of arcs connecting hubs. Moreover, none of them consider other relevant network design decisions such as the activation of other types of arcs or servicing decisions for the O/D nodes.

HNDPPs are also related to other non-competitive network optimization problems, aiming at maximizing the captured demand or optimizing some profit-oriented objective. Examples of the former are the maximal covering location (Church and ReVelle, 1974) or the competitive facility location problem (Aboolian et al., 2007). Examples of the latter are prize-collecting versions of well-known problems that do not consider locational decisions such as traveling salesman (Feillet et al., 2005), vehicle routing (Aras et al., 2011), rural postman (Aráoz et al., 2009), and prize-collecting Steiner tree problems (Álvarez-Miranda et al., 2013), among others.

The above mentioned prize-collecting problems share with HNDPPs a distinguishing feature: they generalize their corresponding *classical* version by incorporating one additional level to the strategic decision-making process, so as to determine the demand customers to be served. In its turn, such decisions induce additional network design decisions. Nevertheless, according to the classification of Contreras and Fernández (2012), all mentioned problems are user-facility demand. That is, service demand relates users (nodes) and service centers (facilities). Instead, HNDPPs are user-user demand, as service demand relates pairs of users among them (O/D nodes of commodities). To the best of our knowledge this is the first time a prize-collecting version of a user-user demand general network design problem is addressed.

3. Formal Definition and Modeling Assumptions of HNDPPs

We can formally define a HNDPP as follows. Let $G = (N, A)$ be a complete directed graph, where $N=\{1, 2, ..., n\}$ represents the set of nodes and A represents the set of arcs. For $(i, j) \in A$, d_{ij} denotes the distance or unit transportation cost between nodes i and j, which we assume to be symmetric, i.e., $d_{ij} = d_{ji}$, and to satisfy the triangle inequality. Let $H \subseteq N$ be the set of potential hub locations and $A_H \subset A$ the subset of arcs connecting two potential hub nodes, i.e. $A_H = \{(i, j) \mid i, j \in H\}$. We also consider the following two sets of undirected edges. The set of edges connecting two potential hubs, denoted as $E_H = \{(i, j) \mid i, j \in H\}$, and the set of edges where at least one endnode is a potential hub, denoted by $E_B = \{(i, j) \mid i \in N, j \in H, i \neq j\}$. For ease of notation, for any edge $(i, j) \in E_H$ we instinctively denote it as (i, j) or (j, i). Since N and H are different sets, for edges in E_B we write $(i, j) \in E_B$, where $i \in N, j \in H$. A *hub edge* $e = (i, j) \in E_H$ connects two different hub nodes i and j and has a per unit flow cost of αd_{ij}. The parameter α $(0 \leq \alpha \leq 1)$ is used as a discount factor to provide reduced unit transportation costs on hub edges to represent economies of scale. Once a hub edge (i, j) is activated, it can be used to route flow in both directions, i.e. from i to j and from j to i. In the literature hub edges are often referred to as hub arcs. Throughout this paper we prefer to maintain the distinction between edges and arcs.

Let K denote the set of commodities. Commodity $k \in K$ is defined as a triplet $(o(k), d(k), W_k)$, where $o(k), d(k) \in N$, respectively denote its origin and its destination, also referred to as its O/D pair, and W_k denotes its service demand, i.e., the amount of flow that must be routed from $o(k)$ to $d(k)$ if commodity k is served. The effect of serving commodity $k \in K$ is threefold. On the one hand it forces the activation of its O/D nodes $o(k)$ and $d(k)$. On the other hand, it produces a per unit profit (prize) P_k, which is independent of the path used to send the commodity demand W_k through the solution network. Finally, serving commodity $k \in K$ also incurs a transportation cost which depends on W_k and on the path that is used to route it from $o(k)$ to $d(k)$. For each $i \in N$, c_i denotes the setup cost for serving node i and for each $i \in H$, f_i is the fixed setup cost for opening a hub at node i. If a node $i \in H$ is selected to be a hub, it is assumed that it will be possible to serve commodities originated (or with destination) at i without activating node i as a servicing node. That is, there is no need to incur in the setup cost c_i for serving node i if it becomes a hub. Each node will thus be exactly one of

the following: a hub node, a served node, or an unserved node.

Similarly to most HLPs, we require all O/D paths to include at least one hub node. That is, the solution network contains no direct connections between two non-hub nodes. Also, as in the case of hub arc location models, we assume that it is always possible to use a so-called *bridge edge* to connect two hub nodes, without benefiting from the discounted unit flow cost of a hub edge (see Campbell et al., 2005). Therefore, the O/D paths associated with served commodities in HNDPPs become more involved than in standard HLPs. In this work we focus on a class of HNDPPs whose solution networks contain at most three edges in each O/D path. In particular, for a given commodity k the path can be represented as $(o(k), i, j, d(k))$ and includes a collection edge from $o(k)$ to hub i, a transfer edge between hubs i and j, and a distribution edge from hub j to $d(k)$. The first and last legs are either *access edges*, connecting a non-hub node to a hub node, or bridge edges, whereas the intermediate leg, if it exists, is a hub edge. That is, some O/D paths consist of just the collection and distribution legs and do not contain a hub edge, i.e. $(o(k), i, i, d(k))$ (origin-hub-destination) with $o(k) \neq i$ and $d(k) \neq i$. Other O/D paths are of the form $(o(k), o(k), d(k), d(k))$ and may arise only when both $o(k)$ and $d(k)$ are hub nodes, consisting of a single hub edge. In paths with at least two edges, if $o(k)$ or $d(k)$ are themselves hub nodes, then the collection or distribution legs also connect two hub nodes but using bridge edges instead. The reader is addressed to Campbell et al. (2005) for an extensive analysis on possibilities for O/D paths in hub location. Taking into account the above mentioned assumptions and requirements on the structure of O/D paths, we define the per unit transportation cost for routing commodity k on the path $(o(k), i, j, d(k))$ as $F_{ijk} = (\chi d_{o(k)i} + \alpha d_{ij} + \delta d_{jd(k)})$, where the parameters χ and δ reflect weights factors for collection and distribution, respectively.

In a HNDPP each edge that is used for sending flow incurs a setup cost. In particular, we denote by r_e the setup cost of hub edge $e \in E_H$. The setup cost of collection/distribution edge is denoted by $q_e, e \in E_B$. This setup cost does not depend on whether they are access or bridge edges. When edges in E_H and E_B are activated, their associated arcs can be used for sending flows in any of their two directions.

The HNDPP consists of (*i*) selecting a set of O/D nodes to be served; (*ii*) locating a set hub facilities; (*iii*) activating a set of hub and collection/distribution edges (access or bridge); (*iv*) selecting a set of commodities to be served, both of whose O/D nodes have been selected in (*i*); and, (*v*)

determining the flows routing the selected commodities through the solution network, with the objective of maximizing the difference between the total profits obtained for routing the service demand of the selected commodities minus the sum of the setup costs for the design of the network and the transportation costs for routing the commodities through the network.

The HNDPP is clearly NP-hard given that it has as a particular case the classical uncapacitated hub location problem with multiple assignments (UHLPMA), which is known to be NP-hard (Contreras, 2015). That is, the HNDPP reduces to the UHLPMA when $s_i = 0$, for $i \in N$, $r_e = 0$, for $e \in E_H$, $q_e = 0$, for $e \in E_B$, and $P_k = \sum_{i \in N} f_i + \max\{F_{ijk} : (i,j) \in A_H\}$, for $k \in K$.

4. Alternative Models and Formulations of HNDPPs

The previous section provides a generic definition of HNDPPs in which their considered strategical and operational decisions are identified. We next propose several alternative HNDPPs which can be applied to a variety of situations. First, we introduce two models which are mainly driven by the total profits obtained from serving a set of commodities. Then, we present another two models which are service-oriented useful to companies with service commitment or with market penetration policies. Finally, we propose two more realistic extensions, which consider that commodities demands are sensitive to price variations. We present Mixed Integer Programming (MIP) formulations for each of these models, which will be computationally tested in Section 5.

Before presenting these models, we define the following sets of decision variables that are common to all of them. For $i \in H$, we introduce binary location variables z_i equal to 1 if and only if a hub is located at node i, and for $i \in N$ we define binary variables s_i equal to 1 if and only if node i is served (i.e. activated as a non-hub node). Additional sets of binary variables are defined to characterize the different types of edges that are used. For $e \in E_H$, we define y_e equal to 1 if and only if hub edge e is activated and for $e \in E_B$, we introduce t_e equal to 1 if and only if edge e is activated as an access or bridge edge. Finally, for $k \in K$, $i, j \in H$, we define routing variables x_{ijk} equal to 1 if and only if commodity k is routed via arc $(i,j) \in A_H$. When $i = j$, $x_{iik} = 1$ represents that commodity k is routed on the path $(o(k), i, d(k))$ visiting only hub i and thus, it is not routed via a hub edge.

4.1. *Profit-oriented Models*

We next present two HNDPPs which are solely driven by profit. In both of them only nodes that increase the system profit will be served. The first one, denoted as PO_1, is more "flexible", in the sense that, among all commodities connecting served O/D nodes, only the ones that are actually profitable will be routed. In PO_1 it is thus possible that a commodity is not routed even if both its origin and destination are activated. It will only be served if routing it produces an additional profit.

As mentioned, such a model can be applicable, for instance, in airline and ground transportation systems. Servicing a city does not mean that connections between this city and any other servicing city in a company's network will be necessarily offered. Only connections between this city and other cities that are profitable will be offered. The PO_1 can be formulated as follows:

$$(PO_1) \quad \text{maximize} \quad \sum_{k\in K}\sum_{(i,j)\in A_H} W_k(P_k - F_{ijk})x_{ijk} - \sum_{i\in H} f_i z_i - \sum_{i\in N} c_i s_i - \sum_{e\in E_H} r_e y_e - \sum_{e\in E_B} q_e l_e \tag{1}$$

$$\text{subject to} \quad \sum_{(i,j)\in A_H} x_{ijk} \le s_{o(k)} + z_{o(k)} \qquad k \in K \tag{2}$$

$$\sum_{(i,j)\in A_H} x_{ijk} \le s_{d(k)} + z_{d(k)} \qquad k \in K \tag{3}$$

$$\sum_{j\in H} x_{ijk} + \sum_{j\in H: i\neq j} x_{jik} \le z_i \qquad k \in K, i \in H \tag{4}$$

$$s_i + z_i \le 1 \qquad i \in H \tag{5}$$

$$x_{ijk} + x_{jik} \le y_e \qquad k \in K, e = (i,j) \in E_H \tag{6}$$

$$\sum_{j\in H} x_{ijk} \le l_{o(k)i} \qquad k \in K, (o(k),i) \in E_B \tag{7}$$

$$\sum_{i\in H} x_{ijk} \le l_{d(k)j} \qquad k \in K, (d(k),j) \in E_B \tag{8}$$

$$x_{ijk} \ge 0 \qquad (i,j) \in A_H, k \in K \tag{9}$$

$$z_i \in \{0,1\} \qquad i \in H \tag{10}$$

$$s_i \in \{0,1\} \qquad i \in N \tag{11}$$
$$y_e \in \{0,1\} \qquad e \in E_H \tag{12}$$
$$t_e \in \{0,1\} \qquad e \in E_B. \tag{13}$$

The first term of the objective function represents the net profit for routing the commodities. The other terms represent the total setup costs of the hubs that are chosen, the non-hub nodes that are selected to be served, and the edges that are used (hub edges and access/bridge edges, respectively). Constraints (2) and (3) impose that the O/D nodes of each routed commodity are activated, either as hub or served nodes. Constraints (4) prevent commodities from being routed via non-hub nodes, whereas constraints (5) forbid that a hub node is also activated as a served node. While constraints (6) activate hub edges, constraints (7) and (8) impose that collection and distribution edges are activated (either as access or bridge edges). Finally, constraints (9) to (13) define the domain for the decision variables. As usual in uncapacitated hub location models, the above formulation does not require to explicitly impose the integrality of the routing variables x. Each commodity, if routed, will use exactly one path of the solution network.

The above formulation has a very large number of variables and constraints. However, we can exploit the following properties to reduce its size.

Property 1. *There is an optimal solution to formulation* (1)–(13) *where* $x_{ijk} = 0$, *for every* $k \in K$ *and* $(i,j) \in A_H$, *with* $P_k - F_{ijk} \leq 0$.

The above property is a direct consequence of the modeling assumption that only profitable commodities will be routed. According to it, for each commodity $k \in K$ all the routing variables whose cost is not strictly smaller that its profit P_k can be eliminated, as routing them will not increase the system profit.

Property 2. *Let* $Q = \{(z,s,y,t,x)$ *that satisfy* $(2)-(13)\}$ *be the domain of feasible solutions to* PO_1. *Then,*

a) For every $k \in K$ *and* $e = (i,j) \in E_H$, $y_e \leq z_i$ *and* $y_e \leq z_j$.
b) For every $i \in H$ *and* $k \in K$, $t_{o(k)i} \leq z_i$ *and* $t_{o(k)i} \leq z_{o(k)} + s_{o(k)}$.

Point a) of Property 2 is a direct consequence of the fact that points (z,s,y,t,x) that satisfy constraints (4) and (6) ensure that $y_e = 1$ if its endnodes are hubs, whereas point b) is a consequence of the fact that any point (z,s,y,t,x) that satisfies constraints (2)–(3) and (7)–(8) has $t_{o(k)i} = 1$

for all i, $o(k)$ such that i is a hub node and $o(k)$ is either a served node or a hub node.

The next service-based model, referred to as PO_2, is more restrictive than PO_1 as it forces to serve any commodity whose O/D nodes are both activated. This constraint can be imposed to the decision maker due to either external regulations or service commitments. An important consequence of this assumption is that the solution network will consist of a single connected component with no isolated hub nodes. PO_2 can be stated as follows:

$$(PO_2) \quad \text{maximize} \quad \sum_{k \in K} \sum_{(i,j) \in A_H} W_k(P_k - F_{ijk}) x_{ijk} - \sum_{i \in H} f_i z_i - \sum_{i \in N} c_i s_i - \sum_{e \in E_H} r_e y_e - \sum_{e \in E_B} q_e l_e$$

$$\text{subject to} \quad (2) - (13)$$

$$s_{o(k)} + z_{o(k)} + s_{d(k)} + z_{d(k)} \leq \sum_{(i,j) \in A_H} x_{ijk} + 1 \quad k \in K. \quad (14)$$

The objective function of PO_2 is the same as that of PO_1. Constraints (14) force commodities to be routed if their O/D nodes are both activated. We note that Property 1 no longer holds for PO_2 because of the addition of constraints (14). As it will be shown in Section 5, this additional requirement considerably increases the complexity for optimally solving PO_2 with a general purpose solver.

4.2. Service-oriented Models

Pure profit-oriented models, such as PO_1 and PO_2, might not be suitable for companies with market penetration policies. In this case, the decision maker might be forced to ensure a predefined presence of a company in the market by servicing a minimum number of customers demands, even if this is suboptimal from a profit perspective. This is the focus of the two models below, where this threshold is imposed through a fraction of the total number of commodities in one case, and through a fraction of the total demand in the other case.

Again, airline transportation can be a field of application for the first model, denoted as SO_1. Governmental or subsidized companies may be requested to provide a country air transportation services between a minimum number of pairs of cities. Applications for the second model, denoted as

SO_2, may arise in postal networks. Depending on their scope, postal delivery companies may be forced by the local or central government to route at least a fraction of the total parcels that need to be routed. The outcome of their decisions on where to locate their hubs and what cities to serve so as to maximize their profits, will indeed depend on this percentage.

Let $0 \leq \beta_1 \leq 1$ denote the minimum fraction of the total number of commodities whose service wants to be guaranteed. Then the formulation of SO_1 that guarantees this service commitment is the following:

$$
\begin{aligned}
(SO_1) \quad \text{maximize} \quad & \sum_{k \in K} \sum_{(i,j) \in A_H} W_k (P_k - F_{ijk}) x_{ijk} - \sum_{i \in H} f_i z_i - \sum_{i \in N} c_i s_i \\
& - \sum_{e \in E_H} r_e y_e - \sum_{e \in E_B} q_e l_e \\
\text{subject to} \quad & (2) - (13) \\
& \sum_{k \in K} \sum_{(i,j) \in A_H} x_{ijk} \geq \beta_1 |K|. \qquad (15)
\end{aligned}
$$

The objective function is the same as in PO_1. Constraints (15) guarantee that among all the commodities with demand, at least a percentage of $100\beta_1$ of them are routed.

In the case of SO_2, which guarantees a minimum service of the total demand, we impose that the overall flow that is routed through the network is at least a fraction $0 \leq \beta_2 \leq 1$ of the overall demand $\sum_{k \in K} W_k$. SO_2 can be stated as follows:

$$
\begin{aligned}
(SO_2) \quad \text{maximize} \quad & \sum_{k \in K} \sum_{(i,j) \in A_H} W_k (P_k - F_{ijk}) x_{ijk} - \sum_{i \in H} f_i z_i - \sum_{i \in N} c_i s_i \\
& - \sum_{e \in E_H} r_e y_e - \sum_{e \in E_B} q_e l_e \\
\text{subject to} \quad & (2) - (13) \\
& \sum_{k \in K} \sum_{(i,j) \in A_H} W_k x_{ijk} \geq \beta_2 \sum_{k \in K} W_k. \qquad (16)
\end{aligned}
$$

The objective function is the same as SO_1. Constraints (16) force the model to route at least a percentage $100\beta_2$ of the total potential flows that can be routed through the network.

4.3. Profit-oriented Models with Multiple Demand Levels

In all previous models, it is assumed that if a commodity $k \in K$ is served then all its demand W_k will be routed and a profit P_k will be perceived. However, in practice, for a given O/D pair the amount of demand W_k that is actually served can be related to the price set to provide such transportation service. That is, the amount of demand that requires service associated with a commodity k will depend on the per unit profit P_k set by the company. Therefore, an additional operational decision can be considered, which is to select for each commodity $k \in K$ the profit level that will allow the company to capture the optimal portion of the total demand W_k.

In this section we consider profit-oriented models with multiple demand levels in which the above mentioned decisions are taken into account. The amount of price-dependent demand that is captured for each commodity, is usually modeled with various nonlinear continuous functions (see, for instance Lüer-Villagra and Marianov, 2013; O'Kelly et al., 2014). In this paper, to keep the model tractable while maintaining the rest of the decisions already considered, we employ a discrete approximation function that considers a set of possible values for commodities demands each of them associated with a profit. That is, we use L as the index set of demand and profit levels for the commodities. For each commodity $k \in K$ and level $l \in L$, let now W_k^l denote the amount of demand that is routed if commodity k is served at level l, and P_k^l the corresponding profit. All other data remains as in the previous models.

To formulate the first profit-oriented model with multiple demand levels, denoted as POM_1, for each $l \in L, i, j \in H$ and $k \in K$, we substitute the original set of routing variables x by an extended set of continuous routing variables, x_{ijk}^l, which denote the fraction of commodity k served at demand level l that is routed via arc $(i, j) \in A_H$. The remaining decision variables are the same as in previous models, since we assume that they do not depend

on demand levels. The POM_1 can be formulated as follows:

$$(POM_1) \text{ maximize} \quad \sum_{l\in L}\sum_{k\in K}\sum_{(i,j)\in A_H} W_k^l(P_k^l - F_{ijk})x_{ijk}^l - \sum_{i\in H} f_i z_i$$
$$- \sum_{i\in N} c_i s_i - \sum_{e\in E_H} r_e y_e - \sum_{e\in E_B} q_e t_e$$

$$\text{subject to} \quad (10) - (13)$$

$$\sum_{l\in L}\sum_{(i,j)\in A_H} x_{ijk}^l \le s_{o(k)} + z_{o(k)} \qquad k \in K \qquad (17)$$

$$\sum_{l\in L}\sum_{(i,j)\in A_H} x_{ijk}^l \le s_{d(k)} + z_{d(k)} \qquad k \in K \qquad (18)$$

$$\sum_{l\in L}\sum_{j\in H} x_{ijk}^l + \sum_{l\in L}\sum_{j\in H: i\neq j} x_{jik}^l \le z_i \quad k \in K, i \in H \qquad (19)$$

$$s_i + z_i \le 1 \qquad i \in H \qquad (20)$$

$$x_{ijk}^l + x_{jik}^l \le y_e \qquad k \in K, e = (i,j) \in E \qquad (21)$$

$$\sum_{l\in L}\sum_{j\in H} x_{ijk}^l \le t_{o(k)i} \quad k \in K, (o(k),i) \in E_B \qquad (22)$$

$$\sum_{l\in L}\sum_{i\in H} x_{ijk}^l \le t_{d(k)j} \quad k \in K, (d(k),j) \in E_B \qquad (23)$$

$$x_{ijk}^l \ge 0 \qquad (i,j) \in A_H, k \in K, l \in L. \qquad (24)$$

The first term of the objective function represents the net profit for routing commodities at their different demand levels. The other terms are as in previous models. Constraints (17)-(24) are the analog to (2)-(9) taking into account the possible demand levels of the commodities.

Given that POM_1 does not consider any capacity constraints on the hubs or edges, it has a very useful property which can be exploited to considerably reduce the size of the above formulation. In particular, it can be shown that there is always an optimal solution to POM_1 in which for each served commodity, exactly one demand level and one path are selected. Moreover, for each commodity $k \in K$, its optimal demand level can be identified a priori. This observation is formalized in the following result.

Proposition 1. *For each $k \in K$, let $\bar{l}_k \in \arg\max_{l\in L}\{W_k^l P_k^l\}$. Then,*

1. *There is an optimal solution to POM_1 where $x_{ijk}^l = 0$, for $l \neq \bar{l}_k$, $(i,j) \in A_H$.*

2. *An optimal solution to POM_1 can found by solving PO_1 with $W_k = W_k^{l_k}$ and $P_k = P_k^{l_k}$, for each $k \in K$.*

Proposition 1 is a direct consequence of the fact that in POM_1 we assume that the demand levels of the commodities have no effect on of the setup costs of the network design decisions, particularly, on the setup costs of the hubs and served nodes.

Instead in the model that we present next, denoted as POM_2, we assume that hubs and served nodes can be activated at different operation levels, incurring setup costs, which depend on the amount of flow that is being processed at the nodes. That is, POM_2 is a capacitated model which considers multiple capacity levels to limit the maximum flow processed at a hub or served node. To this end, we denote as T the index set of operation levels for the hubs and for the served nodes (for ease of notation and without loss of generality we assume they are the same). For each potential hub $i \in H$ and operation level $t \in T$, let f_i^t denote the setup cost for hub i with operation level t, which allows serving a maximum amount of flow φ_i^t. Similarly, for each $i \in N$ and $t \in T$, let c_i^t denote the setup cost for serving node i with operation level t, which allows serving a maximum amount of flow ρ_i^t. We now extend the set of decision variables for the hubs and served nodes to the following. For each $i \in H$ and $t \in T$, variable z_i^t takes the value 1 if and only if a hub is located at node i with operation level t. For $i \in N$ and $t \in T$, variable s_i^t is equal to 1 if and only if node i is served with operation level t. POM_2 can be formulated as follows:

$$
\begin{aligned}
(POM_2)\ \text{maximize} \quad & \sum_{l\in L}\sum_{k\in K}\sum_{(i,j)\in A_H} W_k^l(P_k^l - F_{ijk})x_{ijk}^l - \sum_{i\in H}\sum_{t\in T} f_i^t z_i^t \\
& - \sum_{i\in N}\sum_{t\in T} c_i^t s_i^t - \sum_{e\in E_H} r_e y_e - \sum_{e\in E_B} q_e t_e \\
\text{subject to} \quad & (12)-(13), (21)-(24) \\
& \sum_{l\in L}\sum_{(i,j)\in A_H} x_{ijk}^l \le \sum_{t\in T}(s_{o(k)}^t + z_{o(k)}^t) \qquad k \in K \ (25) \\
& \sum_{l\in L}\sum_{(i,j)\in A_H} x_{ijk}^l \le \sum_{t\in T}\left(s_{d(k)}^t + z_{d(k)}^t\right) \qquad k \in K (26)
\end{aligned}
$$

$$\sum_{l\in L}\left(\sum_{j\in H} x^l_{ijk} + \sum_{j\in H: i\neq j} x^l_{jik}\right) \leq \sum_{t\in T} z^t_i \quad k\in K, i\in H \tag{27}$$

$$\sum_{t\in T} s^t_i + \sum_{t\in T} z^t_i \leq 1 \qquad i\in H \tag{28}$$

$$\sum_{k\in K}\sum_{l\in L} W^l_k\left(\sum_{j\in H} x^l_{ijk} + \sum_{l\in L}\sum_{j\in H: i\neq j} x^l_{jik}\right) \leq \sum_{t\in T}\varphi^t_i z^t_i \qquad i\in H \tag{29}$$

$$\sum_{(i,j)\in A}\sum_{l\in L}\left(\sum_{k\in K: o(k)=h} W^l_k x^l_{ijk} + \sum_{k\in K: d(k)=h} W^l_k x^l_{ijk}\right) \leq \sum_{t\in T}\rho^t_h s^t_h + M\sum_{t\in T} z^t_h \qquad h\in N \tag{30}$$

$$z^t_i, s^t_i \in \{0,1\} \qquad i\in N, t\in T. \tag{31}$$

The objective function and constraints (25)-(28) have a similar interpretation to (1)-(5) of PO_1. Constraints (29) guarantee that the service level at which a hub is open allows to serve all the incoming and outgoing flow that is routed through it. Constraints (30) have a similar interpretation, with respect to the served nodes. They state that the total incoming and outgoing flow at a served node must not exceed its installed operational capacity. The last term $M\sum_{t\in T} z^t_m$ on the right hand side of the constraints is used to deactivate the constraint in case node h becomes a hub node, where M stands for a sufficiently large constant.

Of course more general models could be considered where different operation levels and associated setup costs are considered also for all edges. This will indeed increase further the complexity of the models, although the modeling techniques will be quite similar to the ones we have used so far. We close this section by noting that Property 2 holds for all the considered models.

5. Computational Experiments

We have been run a series of computational experiments to analyze the performance of the different HNDPPs introduced in Section 4. In particular, we study the characteristics of the optimal solutions obtained with each of

the considered models by providing insights into their optimal network structures and by evaluating the effect of their different parameters. We also give numerical results that allow quantifying and comparing the empirical computational difficulty of the different HNDPPs variants and the quality of the associated MIP formulations we have presented. All experiments were run on an HP station with an Intel Xeon CPU E3-1240V2 processor at 3.40 GHz and 24 GB of RAM under Windows 7 environment. All MIP formulations were coded in C and solved using the callback library of CPLEX 12.5.1. We set the search method to be a traditional (deterministic) branch and bound algorithm and we do not allow CPLEX to generate cuts. The rest of CPLEX parameters are set to their default settings.

We have used the well-known CAB data set of the US Civil Aeronautics Board in all the experiments. These instances were obtained from the website *(http://www.researchgate.net/publication/269396247_cab100_mok)*. The data in the CAB set refers to 100 cities in the US. It provides Euclidean distances between cities, d_{ij}, and the values of the service demand between each pair of cities, W_k, where $o(k) \neq d(k)$. We have considered instances with n =15, 20, 25, 30, 40, and 50 and $\alpha \in \{0.2, 0.5, 0.8\}$. Since CAB instances do not provide the setup costs for opening facilities, we use the costs F_i generated by de Camargo et al. (2008). In order to properly adapt these instances for the considered HNDPPs which involve additional setup costs for the design of the hub network, we have made the following modifications to take into account the difference of the magnitudes of the input data. We used a fraction of F_i, as the setup cost of opening hubs, i.e. $f_i = \kappa F_i$, where $\kappa \in [0.35, 2.6]$ is a parameter that depends on the number of nodes n and the discount factor α of each considered instance. The setup costs c_i, $i \in N$, for served nodes are $c_i = \nu f_i$, where $\nu = 0.1$ unless otherwise stated. The setup costs r_e, $e = (i, j) \in E_H$, for activating hub edges are $r_e = \tau(f_i + f_j)/2$, where $\tau \in \{0.2, 0.3, 0.5\}$ is a parameter used to model the increase (decrease) in setup costs on the hub edges when considering smaller (larger) discount factors α. The setup costs q_e, $e = (i, j) \in H_B$, for activating access/bridge edges are set to $q_e = \sigma(f_i + f_j)/2$, where $\sigma = 0.08$ unless otherwise stated. The profits p_k, $k \in K$, for routing commodities are randomly generated as $p_k = \varphi \sum_{(i,j)\in A_H} F_{ijk}/|A_H|$, where φ is a continuous random variable following a uniform distribution $\varphi \sim U\,[0.2, 0.3]$. The collection and distribution factors are $\chi = \delta = 1$.

5.1. Results for Profit-oriented Models

Table 1 summarizes the numerical results obtained for the profit-oriented model PO_1 with a set of 18 instances with up to 50 nodes. The first two columns give information on the instances: n the number of nodes and α the discount factor on hub edges. The next three columns give information about the solution algorithm and its associated bounds. *% LP GAP* shows percentage gaps between the values of the Linear Programming (LP) relaxations and optimal values, computed as $100|v_{LP} - v^*|/v^*$, where v^* and v_{LP} denote the optimal and LP values, respectively. *Time(sec)* give the computing times (in seconds) needed to optimally solve each instance, and *Nodes* the number of nodes explored by CPLEX in the enumeration tree. The last seven columns provide information on the characteristics of optimal solutions. In particular, *Optimal value* give optimal solution values. *Open hubs* and *Served nodes* show the number of open hub facilities and served non-hub nodes, respectively. *Hub edges*, *Access edges*, and *Bridge edges*, give the ratio between the actual number of hub, access and bridge edges and their maximum possible value, respectively. We recall that: *i)* the number of hub edges in a fully interconnected hub-level network is $\sum_{i\in H} z_i(\sum_{i\in H} z_i - 1)/2$, *ii)* the maximum number of access edges that can be used in a hub network, given by a multiple allocation pattern, is $\sum_{i\in N} s_i \sum_{i\in H} z_i$, and *iii)* the maximum number of bridge edges is $\sum_{i\in H} z_i(\sum_{i\in H} z_i - 1)/2$, i.e. one for each candidate hub edge. *% Served O/D pairs* show the percentage of commodities served by the hub network, computed as $100 \sum_{k\in K} \sum_{i,j\in H} x_{ijk}/|K|$. Finally, *% Routed Flows* give the percentage of all the demand that is served, computed as $100 \sum_{k\in K} \sum_{i,j\in H} W_k x_{ijk} / \sum_{k\in K} W_k$.

The results of Table 1 show that CPLEX can solve to optimality all considered instances for the PO_1 with up to 50 nodes in less than 15 minutes. In 8 out of 18 considered instances, the LP relaxation solution is integer. For the remaining instances the % LP gap varies between 0.10 and 1.26, and the number of explored branch and bound nodes is very small (no more than three). The small CPU time required to solve these instances is partially attributed to the effectiveness of Property 1 for eliminating a large number of x_{ijk} variables and constraints (6).

As for the structure of the optimal networks, in all but two instances the hub-level networks are incomplete and the number of hub nodes and hub edges vary from 2 to 6 and from 1 to 9, respectively. In the case of the access edges, in only three instances the activation of these edges corresponds to a single allocation pattern. In the rest of the considered instances, some

n	α	% LP gap	Time (sec)	Nodes	Optimal value	Open Hubs	Served Nodes	Hub Edges	Access Edges	Bridge Edges	%Served O/D pairs	%Routed Flows
15	0.2	0.78	0.45	1	744147.44	5	6	7/10	9/30	1/10	38.57	54.32
15	0.5	0.00	0.23	0	564337.70	3	6	1/3	7/18	0/3	14.76	24.21
15	0.8	1.26	0.25	2	681233.88	3	6	1/3	8/18	0/3	12.86	23.25
20	0.2	0.00	1.42	0	7374782.29	5	9	7/10	10/45	1/10	34.74	66.98
20	0.5	0.00	1.03	0	4444523.79	5	8	4/10	10/40	0/10	21.58	47.78
20	0.8	0.10	1.12	1	3859418.94	3	8	1/3	10/24	0/3	16.58	36.07
25	0.2	0.00	8.56	0	12594354.40	5	15	9/10	16/75	0/10	41.67	71.34
25	0.5	1.08	4.18	3	7782754.18	6	11	6/15	14/66	1/15	23.17	44.73
25	0.8	0.94	3.63	3	7008641.56	5	10	3/10	14/50	1/10	16.33	35.80
30	0.2	0.00	16.86	0	2715686.99	5	17	7/10	17/85	1/10	39.20	65.35
30	0.5	0.37	6.50	0	1628776.61	4	16	3/6	19/64	0/6	23.79	43.99
30	0.8	0.00	4.81	0	1285170.75	3	14	1/3	18/42	0/3	15.75	35.36
40	0.2	0.14	307.30	0	2679979.00	5	22	7/10	22/110	1/10	33.65	61.54
40	0.5	0.11	36.46	0	1593117.21	3	17	3/3	19/51	0/3	21.09	38.31
40	0.8	0.10	20.95	0	1206311.05	2	12	1/1	16/24	0/1	10.32	27.78
50	0.2	0.26	853.71	0	2737146.72	6	26	8/15	26/156	1/15	28.73	60.12
50	0.5	0.00	107.19	0	1659969.31	4	23	3/6	25/92	0/6	17.31	39.60
50	0.8	0.00	55.03	0	1294574.79	3	18	1/3	22/54	0/3	9.31	29.75

Table 1: Computational experiments for PO_1.

served nodes are connected to two or more hubs but no solution provides a (complete) multiple allocation pattern in which each node is connected to every hub. In addition, bridge edges are rarely used. Only one bridge edge is present in the solution network of 7 out of the 18 considered instances. Depending on the size of the instance and the considered discount factor, the served O/D pairs ranges from 9.31% to 41.67%, and the routed flows from 23.25% to 71.34%. This is a direct consequence from the fact that in none of the considered instances, it is optimal to serve all nodes. In fact, the unserved nodes range from 20% to 65%.

In order to evaluate the impact of the setup cost for serving nodes in solution networks, Figure 1 compares the optimal hub networks produced by PO_1 for the CAB instance with $n = 25$ and $\alpha = 0.5$ when setting c_i as 0%, 15% and 40%, of the setup cost f_i. Triangles represent hubs, full circles served nodes, and empty circles unserved nodes. Black lines represent hub edges while gray lines represent access and bridge edges.

When no setup cost for serving nodes is considered (Figure 1a), the optimal solution network consists of two disconnected components with five interconnected hubs, one isolated hub, and 15 served nodes. Note that even though there is no setup cost for serving nodes, there are four unserved nodes. That is, activating the required access edges does not compensate the profits collected from serving a portion of the demand associated with these nodes. In fact, only 29.33% of all O/D pairs are actually served, which in turn ac-

a) $c_i = 0.00\, f_i$
Opt: 8,811,009.83
6 hubs and 15 served nodes
54.38% routed flows

b) $c_i = 0.15\, f_i$
Opt: 7,388,362.45
6 hubs and 10 served nodes
43.57% routed flows

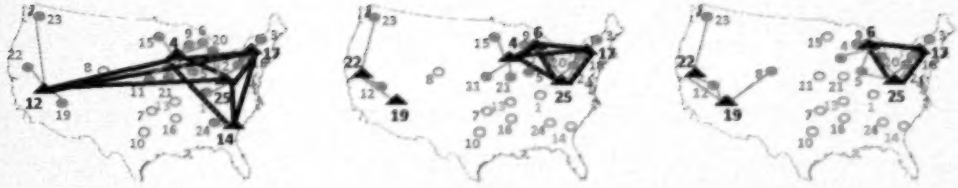
c) $c_i = 0.40\, f_i$
Opt: 5,636,245.08
4 hubs and 8 served nodes
38.80% routed flows

Figure 1: Optimal network for PO_1 with different setup costs c_i with $n = 25$ and $\alpha = 0.5$.

counts for 54.38% of the total amount of flow that can be routed. When increasing the setup costs for serving nodes to $c_i = 0.15 f_i$ (Figure 1b), the number of served nodes is reduced to 10, the number of served O/D pairs decreases to 21.67%, and the routed flows reduces to 43.57%. Note that even though the number of hubs has not changed, three of them change their location and one bridge edge is now used to connect hubs 19 and 22. The total profit is reduced by 16.14%. Finally, when increasing the setup costs to $c_i = 0.40 f_i$ (Figure 1c), the optimal solution network consists of a single connected component having four fully interconnected hubs. The number of served nodes further decreases to eight, the number of served O/D pairs to 20.00% and the routed flows to 38.80%. As a result, the total profit is reduced by 36.03% (with respect to the case in which $c_i = 0$).

We next compare the effect of the discount factor α in solution networks. Figure 2 gives the optimal networks produced by PO_1 for the CAB instance with $n = 25$ and three different values of the discount factor α.

When $\alpha = 0.2$ (Figure 2a), the optimal network consists of a single connected component with five hubs, nine hub edges, 15 served nodes, and 16 access edges. Given the substantial reduction in the transportation costs obtained with such small discount factor, the number of served O/D pairs is 41.67%, which in turn captures 71.34% of the total flow that can be routed. When increasing the discount factor to $\alpha = 0.5$ (Figure 2b), the solution network consists of two disconnected components with one extra hub node but only 11 served nodes. This causes a considerable reduction in both the number of served O/D pairs (21.67%) and in the routed flows (43.57%). Finally, when $\alpha = 0.8$ (Figure 2c), the solution network no longer has node 4 as a



a)α= 0.2
Opt: 12,594,354.40
5 hubs and 15 served nodes
71.34% routed flows

b)α= 0.5
Opt: 7,782,754.18
6 hubs and 11 served nodes
44.73% routed flows

c)α= 0.8
Opt: 7,008,641.56
5 hubs and 10 served nodes
35.80% routed flows

Figure 2: Optimal network for PO_1 with different discount factors α with $n = 25$ and $\nu = 0.10$.

hub and the number of served nodes is decreased by one. As a consequence, the served O/D pairs is further reduced to 16.33% and thus, only 35.80% of the total flow is routed.

We now analyze the results obtained with the second profit oriented model PO_2 in which commodities associated with served O/D nodes are forced to be routed. Table 2 summarizes the numerical results obtained with PO_2 using a set of 15 instances with up to 40 nodes. The columns have the same interpretation as in the previous table.

n	α	*% LP gap*	*Time (sec)*	*Nodes*	*Optimal value*	*Open Hubs*	*Served Nodes*	*Hub Edges*	*Access Edges*	*Bridge Edges*	*%Served O/D pairs*	*%Routed Flows*
15	0.2	0.00	2.17	0	625237.75	5	5	7/10	7/25	1/10	42.86	55.46
15	0.5	0.00	1.69	0	446758.52	2	4	1/1	5/8	0/1	14.29	20.74
15	0.8	5.29	3.90	0	502852.07	2	4	1/1	5/8	0/1	14.29	20.74
20	0.2	0.00	17.71	0	6909102.14	4	7	5/6	8/28	1/6	28.95	54.11
20	0.5	0.00	7.66	0	4214555.69	3	7	3/3	9/21	0/3	23.68	42.40
20	0.8	0.00	7.57	0	3412689.94	2	7	1/1	10/14	0/1	18.95	38.48
25	0.2	0.00	139.93	0	11796725.86	5	12	10/10	13/60	0/10	45.33	69.53
25	0.5	0.01	148.74	0	7375751.43	4	8	6/6	11/32	0/6	22.00	39.90
25	0.8	0.00	39.00	0	6251792.85	3	7	3/3	10/21	0/3	15.00	33.58
30	0.2	0.00	245.37	0	2465434.97	5	17	7/10	20/85	2/10	53.10	71.74
30	0.5	0.52	60.04	0	1592118.18	3	12	3/3	15/36	0/3	24.14	40.26
30	0.8	0.00	45.41	0	1243226.78	2	10	1/1	14/20	0/1	15.17	32.67
40	0.2	0.33	1315.58	0	2452644.20	4	18	6/6	18/72	0/6	29.62	49.55
40	0.5	0.00	267.71	0	1555555.09	3	17	3/3	19/51	0/3	24.36	40.21
40	0.8	0.00	209.45	0	1139492.35	2	12	1/1	16/24	0/1	11.67	29.80

Table 2: Computational experiments for PO_2.

The results of Table 2 show that CPLEX can solve to optimality all considered instances for the PO_2 with up to 40 nodes in less than 22 minutes.

However, preliminary computational experiments showed that CPLEX cannot solve 50 node instances due to memory limitations. In fact, 23GB of memory is not enough to load the problem into CPLEX when considering instances with 50 or more nodes. This is a substantial difference with respect to model PO_1 which only required about 4GB of memory to load a 50 node instance into CPLEX. This difference comes form the fact that Property 1 does not apply to PO_2 and thus, all variables and constraints need to be considered. In 11 out of 15 considered instances, the LP solution is integer and for the remaining instances the % gap varies between 0.01 to 5.29. An interesting observation is that CPLEX can optimally solve all the instances at the root node of the enumeration tree with the pre-processing phase performed after solving the LP relaxation and obtaining optimal solutions with its heuristics.

As for the topology of optimal networks, in all but three instances the solution networks contain between two and five fully interconnected hubs. When comparing the solutions to the corresponding ones with model PO_1, we observe that the number of open hubs and served nodes is always smaller than or equal to that of PO_1. This result was somehow expected as the commodities associated with each served node need to be routed to their (served) destination nodes. In general, the percentage of both served O/D pairs and routed flows slightly decrease with respect to PO_1. Figure 3 compares in detail the optimal solution networks produced by both PO_1 and PO_2 for the CAB instance with $n = 25$ and $\alpha = 0.4$.

a) PO_1
Opt: 7,532,902.74
5 hubs and 13 served nodes
50.87% routed flows

b) PO_2
Opt: 7,107,200.27
3 hubs and 9 served nodes
39.90% routed flows

Figure 3: Optimal networks for PO_1 and PO_2 with $n = 25$ and $\alpha = 0.4$.

On the one side, the solution network of PO_1 (Figure 3a) consists of two disconnected components with four (not fully) interconnected hubs, one isolated hub, and five hub edges. In addition, there are 13 served nodes

connected with 17 access edges and seven unserved nodes. The served O/D pairs is 25.00% and the routed flows are 50.87%. On the other side, the solution network of PO_2 (Figure 3b) consists of a single connected component with only three hubs which are fully interconnected with three hub edges. There are 9 served nodes connected with 13 access edges and 13 unserved nodes. The served O/D pairs are reduced to 22.00% and the routed flows are also reduced to 39.90%. As a result, the total profit decreases by 5.65%.

5.2. Results for Service-oriented Models

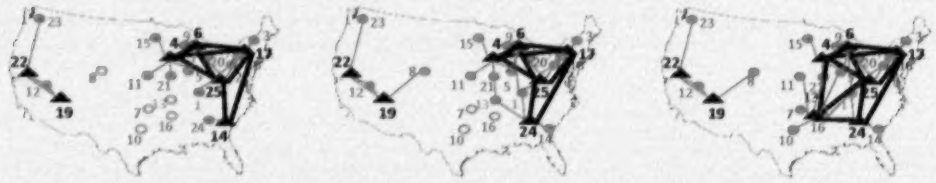
We next analyze the results obtained with the service-oriented models. Table 3 summarizes the numerical results obtained with SO_1 for the same set of instances used for PO_1 and PO_2 with up to 40 nodes. For each considered instance, the value of β_1 was set to represent an increase of 30% with respect to the percentage of served O/D pairs at the optimal solution of PO_1.

$\|N\|$	α	*% LP gap*	*Time(sec)*	*Nodes*	*Optimal*	*Open Hubs*	*Served Nodes*	*Hub Edges*	*Access Edges*	*Bridge Edges*	*%Served O/D pairs*	*%Routed Flows*
15	0.2	1.45	3.08	0	715549.05	6	7	9/15	10/42	1/15	50.48	61.76
15	0.5	3.13	2.27	0	540385.45	4	7	3/6	9/28	0/6	21.43	28.99
15	0.8	2.29	2.36	0	674406.91	3	7	1/3	10/21	0/3	17.14	25.38
20	0.2	0.22	15.69	0	7276304.77	6	11	9/15	12/66	1/15	45.26	75.48
20	0.5	0.03	12.55	0	4343022.03	5	9	4/10	11/45	0/10	28.16	50.38
20	0.8	0.84	14.09	0	3797815.58	4	12	1/6	15/48	0/6	21.58	40.40
25	0.2	0.31	261.74	0	12460393.82	7	16	13/21	18/112	1/21	54.33	79.87
25	0.5	1.51	63.29	0	7725262.89	7	13	8/21	17/91	1/21	30.17	54.36
25	0.8	1.26	39.84	0	6954315.88	5	13	3/10	21/65	1/10	21.33	42.92
30	0.2	0.06	1239.64	0	2670304.47	5	18	7/10	19/90	1/10	51.03	69.89
30	0.5	0.51	647.66	0	1553743.69	4	18	3/6	21/72	0/6	31.03	46.96
30	0.8	1.28	188.57	9	1235795.59	3	16	1/3	20/48	0/3	20.57	38.38
40	0.2	0.32	22452.66	0	2641813.45	5	24	7/10	26/120	1/10	43.78	66.17
40	0.5	0.28	12182.95	0	1523904.76	3	19	3/3	21/57	0/3	27.44	41.46
40	0.8	0.71	1528.43	0	1185550.46	2	16	1/1	20/32	0/1	13.46	30.70

Table 3: Computational Experiments for SO_1.

From Table 3, we note that CPLEX can solve to optimality all considered instances with up to 40 nodes. However, some of these instances take up to six hours to be solved. Contrary to previous models, in none of the considered instances the LP relaxation provides the optimal solution of the integer program. As in PO_2, CPLEX cannot solve 50 node instances due to memory limitations (Property 1 does not hold for SO_1). The % LP gap varies from 0.03 to 3.13 but all instances were solved at root node, except in one case which only required nine nodes. When comparing the structure of optimal networks to that of PO_1, an increase is observed on the number of served nodes (between one to four). This is needed to guarantee the service

constraint on the minimum percentage of served O/D pairs. An interesting observation is that, for the considered instances, an increase of 30% in the fraction of served O/D nodes causes a reduction in the total profit which never exceeds 4.60% and is only 2.35% on average. Figure 4 compares the optimal solution network obtained with SO_1 for the CAB instance with $n = 25$, $\alpha = 0.5$, and increasing values of service parameter β_1.



a) $\beta_1 = 0.3$
Optimal : 7,727,351.75
7 hub nodes and 13 served nodes

b) $\beta_1 = 0.4$
Optimal : 7,433,729.76
7 hub nodes and 15 served nodes

c) $\beta_1 = 0.5$
Optimal : 6,861,175.64
8 hub nodes and 17 served nodes

Figure 4: Optimal networks for SO_1 with different β_1 values with $n = 25$ and $\alpha = 0.5$.

When 30% of the O/D pairs must be served (Figure 4a), the optimal solution consists of two disconnected components with seven hub nodes, eight hub edges, one bridge edge, and 13 served nodes. There are five unserved nodes. When the service requirement is increased to 40% (Figure 4b), the number of hubs remains the same but node 14 is no longer a hub and instead, node 24 becomes a hub. Furthermore, the number of served nodes increases to 15. This 10% increase in the fraction of served O/D pairs produces a total profit decrease of 3.80%. When further increasing the service requirements to 50% (Figure 4c), the optimal solution contains one more hub node and now all nodes are served. This additional 20% increase on the percentage of served O/D pairs, reduces the total profit by 11.21%.

We now analyze the results obtained with the second service oriented model PO_2 in which a service requirement on the minimum percentage of routed flows is considered. Table 4 summarizes the numerical results obtained with SO_2 for the same set of instances used for SO_1. Similar to the previous model, for each considered instance the value of β_2 was set to impose an increase of 30% with respect to the percentage of routed flows at the optimal solution of PO_1.

Table 4 shows that the difficulty for solving SO_2 with CPLEX is slightly less to that of SO_1. All instances can be solved in less than two hours of CPU

$\|N\|$	α	% LP gap	Time(sec)	Nodes	Optimal	Open Hubs	Served Nodes	Hub Edges	Access Edges	Bridge Edges	%Served O/D pairs	%Routed Flows
15	0.2	3.37	5.53	0	631811.05	7	7	10/21	10/49	1/21	58.10	70.66
15	0.5	2.41	5.23	0	522349.50	4	7	3/6	10/28	0/6	25.71	31.55
15	0.8	5.03	2.48	0	644297.69	3	8	1/3	12/24	0/3	22.38	30.36
20	0.2	0.11	28.95	0	6856773.85	7	13	11/21	18/91	2/21	63.16	87.17
20	0.5	0.60	16.81	9	3760447.95	7	13	5/21	19/91	1/21	39.47	62.22
20	0.8	0.10	10.59	0	3577703.13	5	13	3/10	19/65	0/10	28.42	46.92
25	0.2	0.61	690.68	3	11729802.11	8	17	16/28	24/136	2/28	81.83	92.75
25	0.5	1.00	85.29	0	7595790.17	7	14	8/21	20/98	1/21	33.83	58.25
25	0.8	0.36	48.65	0	6817774.63	5	15	3/10	24/75	1/10	25.33	46.70
30	0.2	1.46	1395.73	9	2292482.08	6	21	10/15	25/126	1/15	66.90	84.96
30	0.5	0.49	309.50	3	1366454.09	5	19	4/10	22/95	0/10	34.94	57.20
30	0.8	0.67	215.30	0	1112253.36	3	18	1/3	24/54	0/3	27.13	46.05
40	0.2	0.75	7178.59	0	2359036.65	7	28	11/21	32/196	1/21	57.50	80.01
40	0.5	0.48	1631.89	11	1392785.75	5	23	4/10	25/115	0/10	25.83	49.84
40	0.8	0.11	998.17	0	1128691.80	3	20	1/3	24/60	0/3	15.06	36.31

Table 4: Computational Experiments for SO_2.

time. The % LP gap varies from 0.10 to 5.03 but the number of explored nodes is very small. Similar to model SO_1, in order to guarantee the increased service requirement the number of served nodes increases between one and eight with respect to solution networks of PO_1. The reduction of the total profit caused by an increase of 30% in the fraction of routed flows is now larger (as compared to SO_1), being of 9.76% on the average, and ranging in $[2.40\%, 16.11\%]$. Figure 5 compares the optimal solution network obtained with SO_2 for the CAB instance with $n = 25$, $\alpha = 0.5$, and increasing values of service parameter β_2.

a) β= 0.58
Optimal : 7,595,790.17
7 hubs and 14 served nodes

b) β= 0.67
Optimal : 6,737,205.15
8 hubs and 17 served nodes

c) β= 0.76
Optimal : 5,318,637.67
9 hubs and 16 served nodes

Figure 5: Optimal networks for SO_2 with different β_2 values with $n = 25$ and $\alpha = 0.5$.

When 58% of the total flow is required to be served (Figure 5a), the solution network has two disconnected components with seven hub nodes, eight hub edges, one bridge edge, and 14 served nodes. There are only four unserved nodes. When the service requirement is increased to 67% (Figure

5b), an extra hub node is added to the network (node 13), two hub edges are activated to connect it, and the number of served nodes increases to 17. The total profit decreases by 11.30%. When further increasing the service requirements to 76% (Figure 5c), the optimal solution forms a single connected component containing one more hub (node 10), two extra hub edges, and all nodes are now served. The total profit further decreases by 29.97%. From these analyzes, we note that increasing the minimum service requirements of routed flow seems to have a larger impact on the total profit than increasing the number of served O/D pairs. Figure 6 shows the deterioration of the total profit when increasing the service level requirements in both SO_1 and SO_2 for the CAB instance with $n = 25$, $\alpha \in \{0.2, 0.5, 0.8\}$ and different values of β_1 and β_2, respectively.

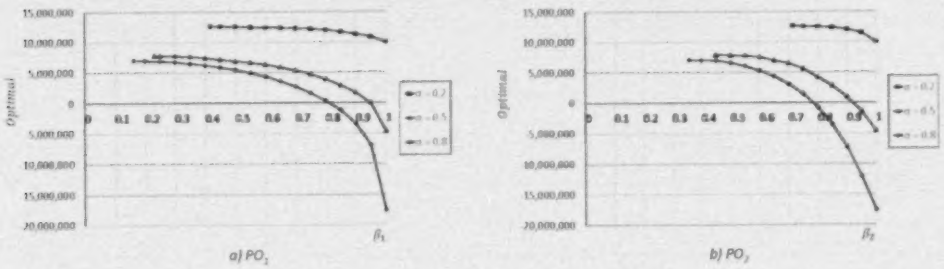


Figure 6: The effect in SO_1 and SO_2 when increasing service levels β_1 and β_2 on the objective value with $N = 25$ and different α values.

Figure 6 shows that the discount factor has a great impact on the deterioration of the total profit in both SO_1 and SO_2 models. In the case of $\alpha = 0.2$, a rather moderate deterioration is perceived when $\beta_1 \geq 0.41$ and $\beta_2 \geq 0.71$ for SO_1 and SO_2, respectively. In fact, when all O/D pairs are required to be served ($\beta_1 = 1$), or alternatively, all flow needs to be routed ($\beta_2 = 1$), the total profit only decreases by 20.57% with respect to model PO_1. However, in the case of $\alpha = 0.5$ and $\alpha = 0.8$, the deterioration of the total profit starts sooner and is much more pronounced. For SO_1, a deterioration is perceived when $\beta_1 \geq 0.23$ for $\alpha = 0.5$, and $\beta_1 \geq 0.16$ for $\alpha = 0.8$. Losses (negative profits) start to appear when $\beta_1 \geq 0.94$ for $\alpha = 0.5$, and $\beta_1 \geq 0.81$ for $\alpha = 0.8$. For SO_2, a reduction of the total profits arises when $\beta_2 \geq 0.45$ for $\alpha = 0.5$, and $\beta_2 \geq 0.36$ for $\alpha = 0.8$. Negative profits occur when $\beta_2 \geq 0.92$ for $\alpha = 0.5$, and $\beta_2 \geq 0.78$ for $\alpha = 0.8$.

5.3. Results for Profit-oriented Models with Multiple Demand Levels

In this last part of the computational experiments we analyze the results obtained with the profit-oriented models with multiple demand levels introduced in Section 4.3. We have adapted the CAB instances used in the previous sections to incorporate the set of demand and profit levels for the commodities and the multiple capacity levels for the hub facilities in the following way. For each $k \in K$, we set W_k^1 and P_k^1 to W_k and P_k, respectively. Data for the other levels are generated by decreasing demand and increasing profit. That is, we defined $W_k^l = W_k^{l-1} \times 0.3$ and $P_k^l = P_k^{l-1} \times 1.2$ for $l = 2, \ldots, |L|$. In addition, for each $i \in H$, we set $f_i^1 = f_i$ and $f_i^t = 0.9 \times f_i^{t-1}$, $t = 2, ..., |T|$. We generated in the same way different levels of setup costs for served nodes. That is, for each $i \in N$, $c_i^1 = c_i$ and $c_i^t = 0.9 \times c_i^{t-1}$, $t = 2, ..., |T|$. For model POM_2, we have also generated different levels of capacities for the hub and served nodes. For each $i \in H$, we set $\varphi_i^1 = \lambda \sum_{i \in H} O_i / \sum_{i \in H} z_i^*$, where λ is a continuous random variable following a uniform distribution $\lambda \sim U[0.9, 1.1]$, and O_i is the total flow passing through hub i at the optimal solution of POM_1 (denoted as z^*). For other capacity levels of hub nodes, $\varphi_i^t = 0.7 \times \varphi_i^{t-1}$, $t = 2, ..., |T|$. Capacity of the served nodes are generated as a fraction of the capacities for the hubs, i.e. $\rho_i^t = \gamma \times \varphi_i^t$, $t = 1, ..., |T|$, and $\gamma = 0.5$. Finally, for these experiments we have considered $|L| = |T| = 5$. Tables 5 and 6 summarize the numerical results obtained for models POM_1 and POM_2, respectively.

$\|N\|$	α	% LP gap	Time(sec)	Nodes	Optimal	Open Hubs	Served Nodes	Hub Edges	Access Edges	Bridge Edges	%Served O/D pairs	%Routed Flows
15	0.2	0.70	0.42	0	777208.91	5	7	7/10	10/35	1/10	53.33	25.35
15	0.5	0.40	0.29	0	568784.33	3	6	1/3	7/18	0/3	19.52	9.25
15	0.8	2.32	0.44	5	703213.17	3	8	1/3	12/24	0/3	28.10	17.34
20	0.2	0.00	2.06	0	7396803.49	5	9	7/10	10/45	1/10	42.37	31.54
20	0.5	0.00	1.00	0	4510122.40	5	9	4/10	11/45	0/10	26.84	24.26
20	0.8	0.18	0.86	0	3038893.43	3	9	1/3	11/27	0/3	20.00	22.03
25	0.2	0.00	29.72	0	12799177.47	5	15	9/10	16/75	0/10	54.17	32.19
25	0.5	1.13	6.14	5	7904939.37	6	14	8/15	16/84	0/15	35.67	26.48
25	0.8	0.90	2.91	0	7183091.68	5	11	3/10	15/55	1/10	20.83	25.56
30	0.2	0.00	40.09	0	2744947.17	5	18	7/10	18/90	1/10	49.31	35.25
30	0.5	0.32	10.95	0	1638012.94	4	16	3/6	19/64	0/6	26.44	26.88
30	0.8	0.27	7.13	0	1298075.33	3	14	1/3	18/42	0/3	17.47	28.78
40	0.2	0.19	793.47	0	2714501.88	5	22	7/10	22/110	1/10	38.78	8.73
40	0.5	0.10	103.24	0	1608061.56	3	17	3/3	19/51	0/3	24.17	4.85
40	0.8	0.41	41.32	0	1220820.46	2	15	1/1	18/30	0/1	15.06	8.11
50	0.2	0.28	1975.82	3	2770380.38	6	26	8/15	26/156	1/15	32.16	8.55
50	0.5	0.07	242.87	0	1675276.99	4	23	3/6	25/92	0/6	19.51	16.80
50	0.8	0.00	93.69	0	1306843.70	3	18	1/3	22/54	0/3	10.57	19.72

Table 5: Computational Experiments for POM_1.

Table 5 shows that CPLEX can solve to optimality all considered instances for the POM_1 with up to 50 nodes in less than 35 minutes. These instances can be efficiently solved by CPLEX mainly due to the fact that Proposition 1 allow us to transform any instance of the POM_1 to an equivalent instance of the PO_1, which in turn, benefits form Property 1 to remove a considerable amount of variables and constraints. When comparing the structure of optimal networks to that of PO_1, a slight increase is observed on the number of served nodes (between one to three) in eight out of the 18 considered instances. In addition, the incorporation of multiple demand and profit levels for the commodities allow the total profit to increase by 1.46% on the average, with a range of $[0.30\%, 4.44\%]$.

$\|N\|$	α	*% LP gap*	*Time(sec)*	*Nodes*	*Optimal*	*Open Hubs*	*Served Nodes*	*Hub Edges*	*Access Edges*	*Bridge Edges*	*%Served O/D pairs*	*%Routed Flows*
15	0.2	12.40	77.55	578	962975.71	6	6	8/15	9/36	1/15	44.76	55.24
15	0.5	7.92	10.51	219	804995.58	6	6	5/15	9/36	1/15	38.57	32.34
15	0.8	9.48	4.90	219	889772.53	4	8	2/6	13/32	1/6	51.43	29.76
20	0.2	6.61	1168.23	1641	6531689.76	7	10	9/21	14/70	2/21	56.84	71.12
20	0.5	6.12	228.42	1137	4501834.83	5	11	3/10	15/55	0/10	27.37	41.34
20	0.8	3.21	32.32	298	4260302.63	5	11	2/10	16/55	1/10	26.84	41.13
25	0.2	6.28	83212.94	13620	11951526.77	10	14	14/45	21/140	5/45	77.17	76.29
25	0.5	4.13	2377.53	1094	8414559.54	7	14	5/21	20/98	3/21	37.33	48.55
25	0.8	3.45	439.39	1022	8093763.96	7	15	3/21	22/105	2/21	27.83	39.46
30	0.2	9.16*	time	2496	1353291.73*	-	-	-	-	-	-	-
30	0.5	15.61	11573.70	2888	1060538.02	5	13	3/10	16/65	1/10	18.85	34.35
30	0.8	17.73	3357.04	2434	988242.85	3	12	0/3	14/36	1/3	12.53	25.69

Table 6: Computational Experiments for POM_2 with $|T| = 5$.

As for POM_2, the results of Table 6 show that it is considerably more difficult to solve. CPLEX can solve to optimality 11 out of the 12 considered instances in one day of CPU time. For one instance ($n = 30$ and $\alpha = 0.2$), the remaining optimality gap was still 9.00% after the CPU time limit was reached. The incorporation of capacity constraints at the hub and served nodes substantially deteriorates the quality of the LP bounds. The % LP gap now varies between 3.21 and 17.73. When comparing the topology of optimal networks to that of POM_1, an increase on the number of hub nodes is observed (between one to five), but the number of served nodes increases or decreases depending on the considered instance. As for the objective value, in one half of the instances the total profit increases (between 6.45% to 41.53%) and in the other half it decreases (between 0.20% to 50.70%).

6. Conclusions

In this paper we introduced a class of hub network design problems with a profit-oriented objective. These problems integrate several locational and network design decisions such as the selection of origin/destination nodes, a set of commodities to serve, and a set of access, bridge and hub edges. They consider the simultaneous optimization of the collected profit, the setup costs of the hub network and the total transportation cost. We proposed three alternative types of models: pure profit-oriented, service-oriented, and profit-oriented with multiple demand levels. Each model was analyzed and a mathematical programming formulation was computationally tested using a general purpose solver. Given the inherent difficulty of the considered models, CPLEX was only able to solve small to medium-size problems. The authors are currently working on the development of ad-hoc decomposition techniques to efficiently solve more realistic, large-scale instances for this challenging class of hub location problems.

Acknowledgments

This research was partially funded by the Spanish Ministry of Science and Education [Grant MTM2012-36163-C06-05 and ERDF funds]. The research of the first two authors was partially funded by the Canadian Natural Sciences and Engineering Research Council [Grant 418609-2012]. This support is gratefully acknowledged.

References

Aboolian, R., O. Berman, and D. Krass (2007). Competitive facility location and design problem. *European Journal of Operational Research 182*(1), 40–62.

Adler, N. and K. Smilowitz (2007). Hub-and-spoke network alliances and mergers: Price-location competition in the airline industry. *Transportation Research Part B 41*(4), 394–409.

Alumur, S. and B. Y. Kara (2008). Network hub location problems: The state of the art. *European Journal of Operational Research 190*(1), 1–21.

Álvarez-Miranda, E., I. Ljubić, and P. Toth (2013). Exact approaches for solving robust prize-collecting steiner tree problems. *European Journal of Operational Research 229*(3), 599–612.

Aráoz, J., E. Fernández, and O. Meza (2009). An LP based algorithm for the privatized rural postman problem. *European Journal of Operational Research 196*(5), 866–896.

Aras, N., D. Aksen, and M. Tuğrul Tekin (2011). Selective multi-depot vehicle routing problem with pricing. *Transportation Research Part C 19*(5), 866–884.

Asgari, N., R. Z. Farahani, and M. Goh (2013). Network design approach for hub ports-shipping companies competition and cooperation. *Transportation Research Part A 48*, 1–18.

Campbell, J. F. (1994). Integer programming formulations of discrete hub location problems. *European Journal of Operational Research 72*(2), 387–405.

Campbell, J. F. (2013). A continuous approximation model for time definite many-to-many transportation. *Transportation Research Part B 54*, 100–112.

Campbell, J. F., A. Ernst, and M. Krishnamoorthy (2005). Hub arc location problems part I: Introduction and results. *Management Science 51*(10), 1540–1555.

Campbell, J. F. and M. E. O'Kelly (2012). Twenty-five years of hub location research. *Transportation Science 46*(2), 153–169.

Church, R. and C. R. ReVelle (1974). The maximal covering location problem. *Papers in Regional Science 32*(1), 101–118.

Contreras, I. (2015). Hub location problems. In G. Laporte, F. Saldanha da Gama, and S. Nickel (Eds.), *Location Science*, pp. 311–344. Springer.

Contreras, I. and E. Fernández (2012). General network design: A unified view of combined location and network design problems. *European Journal of Operational Research 219*(3), 680–697.

Contreras, I. and E. Fernández (2014). Hub location as the minimization of a supermodular set function. *Operations Research 62*(3), 557–570.

Contreras, I., E. Fernández, and A. Marín (2010). The tree of hubs location problem. *European Journal of Operational Research 202*(2), 390–400.

Contreras, I., M. Tanash, and N. Vidyarthi (2015). A branch and cut algorithm for the cycle hub location problem. *submitted*.

de Camargo, R. S., G. Miranda, and H. P. Luna (2008). Benders decomposition for the uncapacitated multiple allocation hub location problem. *Computers and Operations Research 35*(4), 1047–1064.

Eiselt, H. A. and V. Marianov (2009). A conditional p-hub location problem with attraction functions. *Computers and Operations Research 36*(12), 3128–3135.

Feillet, D., P. Dejax, and M. Gendreau (2005). Traveling salesman problems with profits. *Transportation Science 39*(2), 188–205.

Gelareh, S., S. Nickel, and D. Pisinger (2010). Liner shipping hub network design in a competitive environment. *Transportation Research Part E 46*(6), 991–1004.

Hwang, Y. H. and Y. H. Lee (2012). Uncapacitated single allocation p-hub maximal covering problem. *Computers and Industrial Engineering 63*(2), 382–389.

Labbé, M. and H. Yaman (2008). Solving the hub location problem in a star–star network. *Networks 51*(1), 19–33.

Lin, C.-C. and S.-C. Lee (2010). The competition game on hub network design. *Transportation Research Part B 44*(4), 618–629.

Lowe, T. and T. Sim (2012). The hub covering flow problem. *Journal of the Operational Research Society 64*(7), 973–981.

Lüer-Villagra, A. and V. Marianov (2013). A competitive hub location and pricing problem. *European Journal of Operational Research 231*(3), 734–744.

Marianov, V., D. Serra, and C. ReVelle (1999). Location of hubs in a competitive environment. *European Journal of Operational Research 114*(2), 363–371.

Martins de Sá, E., I. Contreras, and J.-F. Cordeau (2015). Exact and heuristic algorithms for the design of hub networks with multiple lines. *European Journal of Operational Research*. DOI:10.1016/j.ejor.2015.04.017.

Martins de Sá, E., I. Contreras, J.-F. Cordeau, R. S. de Camargo, and G. de Miranda (2015). The hub line location problem. *Transportation Science*. DOI:10.1287/trsc.2014.0576.

O'Kelly, M., H. Luna, R. S. de Camargo, and G. de Miranda (2014). Hub location problems with price sensitive demands. *Network and Spatial Economics*. DOI:10.1007/s11067-014-9276-0.

Sasaki, M., J. F. Campbell, A. Ernst, and M. Krishnamoorthy (2014). A Stackelberg hub arc location model for a competitive environment. *Computers and Operations Research 47*, 27–41.

Yildiz, B. and O. Karasan (2015). Regenerator location problem and survivable extensions: A hub covering location persepctive. *Transportation Research Part B 71*, 32–55.